Proofs Homework Set 13

MATH 217 — WINTER 2011

Due April 13

PROBLEM 13.1. Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for $\mathbb{R}^n$. Recall from the previous assignment that we defined the inner product $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}}$.

(a) Find a matrix M such that $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{B}} = \mathbf{u}^T M \mathbf{v}$.

Proof. Let $P = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$. This invertible matrix satisfies the equation $\mathbf{v} = P[\mathbf{v}]_{\mathcal{B}}$ or, equivalently, $[\mathbf{v}]_{\mathcal{B}} = P^{-1}\mathbf{v}$ for every vector $\mathbf{v} \in \mathbb{R}^n$. By definition of $\langle \bullet, \bullet \rangle_{\mathcal{B}}$, we then have that

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{B}} &= [\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}} = (P^{-1}\mathbf{u}) \cdot (P^{-1}\mathbf{v}) \\ &= (P^{-1}\mathbf{u})^T (P^{-1}\mathbf{v}) = \mathbf{u}^T (P^{-1})^T P^{-1} \mathbf{v}. \end{aligned}$$

Therefore the matrix $M = (P^{-1})^T P^{-1}$ works as directed. □

(b) Show that if $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is an orthonormal basis, then $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{B}} = \mathbf{u} \cdot \mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Proof. In this case, the matrix $U = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ has orthonormal columns. Therefore $U^T U = I$ by Theorem 6.6. By the Invertible Matrix Theorem, it follows that $U^{-1} = U^T$. The formula for the matrix M of part (a) then becomes

$$M = (U^{-1})^T U^{-1} = (U^T)^T U^T = U U^T = U U^{-1} = I.$$

Thus, $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{B}} = \mathbf{u}^T \mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. □

(c) Find an example to show that part (b) is not necessarily true if $\mathcal{B}$ is not an orthonormal basis.

Proof. Consider the basis $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for $\mathbb{R}^2$. The $\mathcal{B}$-coordinates of these basis vectors are $[\mathbf{b}_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $[\mathbf{b}_2]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Therefore $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle_{\mathcal{B}} = [\mathbf{b}_1]_{\mathcal{B}} \cdot [\mathbf{b}_2]_{\mathcal{B}} = 0$, but $\mathbf{b}_1 \cdot \mathbf{b}_2 = 1$. □

(d) Suppose now that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Prove that the i, j entry of $[T]_{\mathcal{B}}$ is $\langle T(\mathbf{b}_j), \mathbf{b}_i \rangle_{\mathcal{B}}$.

Proof. The $\mathcal{B}$-matrix of T is given by the formula

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{B}} & [T(\mathbf{b}_2)]_{\mathcal{B}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{B}} \end{bmatrix}.$$

By Theorem 6.5, we may compute the $\mathcal{B}$-coordinates of $T(\mathbf{b}_j)$ via the formula

$$T(\mathbf{b}_j) = (T(\mathbf{b}_j) \cdot \mathbf{b}_1)\mathbf{b}_1 + (T(\mathbf{b}_j) \cdot \mathbf{b}_2)\mathbf{b}_2 + \cdots + (T(\mathbf{b}_j) \cdot \mathbf{b}_n)\mathbf{b}_n.$$

(This formula is a little simpler than that of Theorem 6.5 because each $\mathbf{b}_i$ is a unit vector, which means that $\mathbf{b}_i \cdot \mathbf{b}_i = 1$.) Therefore, we have

$$[T(\mathbf{b}_j)]_{\mathcal{B}} = \begin{bmatrix} T(\mathbf{b}_j) \cdot \mathbf{b}_1 \\ T(\mathbf{b}_j) \cdot \mathbf{b}_2 \\ \vdots \\ T(\mathbf{b}_j) \cdot \mathbf{b}_n \end{bmatrix}.$$

It follows that the i, j entry of $[T]_{\mathcal{B}}$ is $T(\mathbf{b}_j) \cdot \mathbf{b}_i$, which equals $\langle T(\mathbf{b}_j), \mathbf{b}_i \rangle_{\mathcal{B}}$ by part (b). □

PROBLEM 13.2. Let W be a subspace of $\mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation given by $T(\mathbf{x}) = \mathrm{proj}_W(\mathbf{x})$.

(a) Show that for every $\mathbf{x} \in \mathbb{R}^n$, $||T(\mathbf{x})|| \leq ||\mathbf{x}||$.

Proof. By Theorem 6.8, we may write

$$\mathbf{x} = T(\mathbf{x}) + \mathbf{z},$$

where $T(\mathbf{x}) \in W$ and $\mathbf{z} \in W^{\perp}$. Note that this means that $T(\mathbf{x}) \cdot \mathbf{z} = 0$. Therefore

$$\begin{aligned} ||\mathbf{x}||^2 &= \mathbf{x} \cdot \mathbf{x} = (T(\mathbf{x}) + \mathbf{z}) \cdot (T(\mathbf{x}) + \mathbf{z}) \\ &= T(\mathbf{x}) \cdot T(\mathbf{x}) + 2(T(\mathbf{x}) \cdot \mathbf{z}) + \mathbf{z} \cdot \mathbf{z} \\ &= ||T(\mathbf{x})||^2 + ||\mathbf{z}||^2. \end{aligned}$$

Since $||\mathbf{z}||^2 \geq 0$, we conclude that $||\mathbf{x}||^2 \geq ||T(\mathbf{x})||^2$. This last inequality is equivalent to $||\mathbf{x}|| \geq ||T(\mathbf{x})||$ since lengths are nonnegative quantities. □

(b) Show that for every $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \cdot T(\mathbf{x}) \geq 0$.

Proof. As in part (a), let us write $\mathbf{x} = T(\mathbf{x}) + \mathbf{z}$ where $\mathbf{z} \in W^{\perp}$. Then

$$\mathbf{x} \cdot T(\mathbf{x}) = (T(\mathbf{x}) + \mathbf{z}) \cdot T(\mathbf{x}) = T(\mathbf{x}) \cdot T(\mathbf{x}) + \mathbf{z} \cdot T(\mathbf{x}) = T(\mathbf{x}) \cdot T(\mathbf{x})$$

since $\mathbf{z} \cdot T(\mathbf{x}) = 0$. Since $T(\mathbf{x}) \cdot T(\mathbf{x}) \geq 0$ by Theorem 6.1(d), it follows that $\mathbf{x} \cdot T(\mathbf{x}) \geq 0$. □

(c) Define $S : \mathbb{R}^n \to \mathbb{R}^n$ by $S(\mathbf{x}) = \mathbf{x} - T(\mathbf{x})$. Show that this is the orthogonal projection onto $W^{\perp}$.

Proof. By Theorem 6.8, we have a *unique* decomposition

$$\mathbf{x} = \operatorname{proj}_W(\mathbf{x}) + \mathbf{z} = T(\mathbf{x}) + S(\mathbf{x})$$

where $T(\mathbf{x}) = \operatorname{proj}_W(\mathbf{x}) \in W$ and $S(\mathbf{x}) = \mathbf{z} \in W^\perp$. Note that $T(\mathbf{x}) \in (W^\perp)^\perp$ since $T(\mathbf{x}) \cdot \mathbf{v} = 0$ for every $\mathbf{v} \in W^\perp$. (In fact, it is not difficult to show that $(W^\perp)^\perp = W$.)

By Theorem 6.8 again, we have a *unique* decomposition

$$\mathbf{x} = \operatorname{proj}_{W^\perp}(\mathbf{x}) + \mathbf{z}'$$

where $\operatorname{proj}_{W^\perp}(\mathbf{x}) \in W^\perp$ and $\mathbf{z}' \in (W^\perp)^\perp$. On the other hand, we have

$$\mathbf{x} = S(\mathbf{x}) + T(\mathbf{x})$$

where $S(\mathbf{x}) \in W^\perp$ and $T(\mathbf{x}) \in (W^\perp)^\perp$. By the *uniqueness* part of Theorem 6.8, we must then have $\operatorname{proj}_{W^\perp}(\mathbf{x}) = S(\mathbf{x})$ and $\mathbf{z}' = T(\mathbf{x})$.

Therefore, we have shown that $T(\mathbf{x}) = \operatorname{proj}_W(\mathbf{x})$ and $S(\mathbf{x}) = \operatorname{proj}_{W^\perp}(\mathbf{x})$. □

(d) Show that $||\mathbf{x}||^2 = ||T(\mathbf{x})||^2 + ||S(\mathbf{x})||^2$.

Proof. By part (c), we know that $T(\mathbf{x}) = \operatorname{proj}_W(\mathbf{x}) \in W$ and that $S(\mathbf{x}) = \operatorname{proj}_{W^\perp}(\mathbf{x}) \in W^\perp$. It follows that $T(\mathbf{x}) \cdot S(\mathbf{x}) = 0$. Therefore, since $\mathbf{x} = T(\mathbf{x}) + S(\mathbf{x})$, we see that

$$\begin{aligned} ||\mathbf{x}||^2 &= \mathbf{x} \cdot \mathbf{x} = (T(\mathbf{x}) + S(\mathbf{x})) \cdot (T(\mathbf{x}) + S(\mathbf{x})) \\ &= T(\mathbf{x}) \cdot T(\mathbf{x}) + 2(T(\mathbf{x}) \cdot S(\mathbf{x})) + S(\mathbf{x}) \cdot S(\mathbf{x}) \\ &= ||T(\mathbf{x})||^2 + ||S(\mathbf{x})||^2. \end{aligned}$$

□